

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 1, 2003, 15:21:24 ; Search time 2812.92 Seconds  
(without alignments)  
12850.800 Million cell updates/sec

Title: US-09-625-573-1  
Perfect score: 2232  
Sequence: 1 GGATTGAACAGGACGCATT.....TATAACTATGTTGATAAAG 2232

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estmu:\*
  - 5: em\_estov:\*
  - 6: em\_estpl:\*
  - 7: em\_estro:\*
  - 8: em\_hic:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_hic:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: gb\_gss:\*
  - 18: em\_gss\_hum:\*
  - 19: em\_gss\_inv:\*
  - 20: em\_gss\_pin:\*
  - 21: em\_gss\_vrt:\*
  - 22: em\_gss\_fun:\*
  - 23: em\_gss\_mam:\*
  - 24: em\_gss\_mus:\*
  - 25: em\_gss\_other:\*
  - 26: em\_gss\_pro:\*
  - 27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	459.8	21.0	807	13	BI764263 603045953
2	456.6	20.5	469	10	AV715904 AV715904
3	442.8	19.8	457	9	AA694175 AA694175
4	421.4	18.9	563	9	AA034153 AA034153
5	394.8	17.7	398	9	AA807940 AA807940
6	390.2	17.5	407	14	BQ027284 UI-H-COO-

c	7	360	16.1	467	14	H58254
c	8	357.6	16.0	487	10	BB656336
	9	346.8	15.5	461	14	H58245
	10	326.2	14.6	542	9	AA547303
	11	302	13.5	410	14	H58584
	12	296.4	13.3	480	10	AA657263
	13	293.4	13.1	422	14	H58597
	14	292.6	13.1	447	9	AA671573
	15	289	12.9	1074	14	BM917063
c	16	268.2	12.0	289	9	AA034154
	17	266.6	11.9	907	9	AL552677
	18	249.8	11.2	789	12	BG205056
	19	236.4	10.6	876	13	BI906283
	20	235.2	10.5	884	12	BG182330
	21	220.8	9.9	952	14	BM917763
	22	216.8	9.7	745	12	BG204024
c	23	214.6	9.6	1167	14	BQ053936
	24	212.8	9.5	452	9	AI851510
	25	197	8.8	672	10	BB638766
	26	189	8.5	780	13	BI908127
	27	188.4	8.4	568	12	BF193051
	28	186.8	8.4	551	12	BF081124
	29	181	8.1	863	13	BI661279
	30	180.6	8.0	669	13	BI393893
	31	179.4	8.0	533	12	BF193021
	32	177.2	7.9	639	10	BB629533
	33	175.2	7.8	723	14	BM951933
	34	173.6	7.8	965	14	BQ944555
	35	172.2	7.7	934	12	BG460984
	36	171	7.7	207	12	BG315432
	37	169.2	7.6	787	9	AU080004
	38	164.4	7.4	3005	11	AK019478
	39	163.8	7.3	936	12	BF119806
	40	153.2	6.9	665	10	BB628567
	41	151.8	6.8	662	10	BB644135
	42	151.6	6.8	600	13	BI987229
	43	151.6	6.8	646	10	BB615654
	44	147.4	6.6	869	12	BF119225
	45	147.2	6.6	867	13	BI106339

## ALIGNMENTS

RESULT 1	BI764263	807 bp	mrna	linear	EST 25-SEP-2001
LOCUS	603045953F1 NIH_MGC_116	Homo sapiens	cdna	clone	IMAGE:5186388 5',
DEFINITION	mRNA sequence.				
ACCESSION	BI764263				
VERSION	BI764263.1	GI:15755841			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 807)				
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a> Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM11465 row: i column: 13 High quality sequence stop: 805. Location/Qualifiers 1. .807				
FEATURES	source				

H58254 YR06H06.s1  
BB656336 UI-M-BHO-  
H58245 YR06F06.s1  
AA547303 YK28H04.r  
H58584 YR06F06.r1  
AA657263 YR06H06.r1  
H58597 YR06H06.r1  
AA671573 YR06H06.r1  
BM917063 AGENCOURT  
AA034154 Z106F10.s  
AL552677 AL552677  
BG205056 RST24475  
BI906283 603063222  
BG182330 RST1196 A  
BM917763 AGENCOURT  
BG204024 RST23417  
BQ053936 AGENCOURT  
AI851510 UI-M-BHO-  
BB638766 BB638766  
BI908127 603067423  
BF193051 244485 MA  
BF081124 233997 MA  
BI661279 603304243  
BI393893 P9P1n.pk0  
BF193021 244450 MA  
BB629533 BB629533  
BM951933 UI-M-EHO-  
BQ944555 AGENCOURT  
BG460984 RST43688  
BG315432 P02.0.308  
AU080004 AU080004  
AK019478 Mus muscu  
BF119806 601758673  
BB628567 BB628567  
BB644135 BB644135  
BI987229 3193-25 M  
BB615654 BB615654  
BF119225 601737232  
BI106339 602890674



```
QY 1655 GCTGTCATCTCAGCTGGATCTCCATTCCTCAGGCTTGCTGCC 1697
|||||
Db 421 GCTGTCATCTCAGCTGGATCTCCATTCCTCAGGCTTGCTGCC 463

RESULT 3
AA694175/c 457 bp mRNA linear EST 16-DEC-1997
LOCUS z142a04.s1 Soares_fetal_liver_spleen_INFLS_S1 Homo sapiens cDNA
DEFINITION clone IMAGE:433422 3', mRNA sequence.
ACCESSION AA694175
VERSION AA694175.1 GI:2695113
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 457)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST project
Unpublished (1997)
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 455.

FEATURES
source
1..457
/organism="Homo sapiens"
/db_xref="GDB:1333191"
/db_xref="taxon:9606"
/clone="IMAGE:433422"
/clone_lib="Soares_fetal_liver_spleen_INFLS_S1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5'
AACTGGAAGAATTAATAAGATCTTTTGTGTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 137 a 96 c 82 g 142 t
ORIGIN
Query Match 19.8%; Score 442.8; DB 9; Length 457;
Best Local Similarity 99.3%; Pred. No. 6.3e-104;
Matches 455; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1770 TCGCAGATGCTTGATCTCATATTTGTCCTATTTTCCAGTGGGAACTCCTAAATCA 1829
|||||
Db 457 TCGCAGATGCTTGATCTCATATTTGTCCTATTTTCCAGTGGGAACTCCTAAATCA 399
|||||
QY 1830 AATTGGCTTCTAATCAAGCTTTTAAACCCCTATTGCTTAAGAAATGGAAGGTGGAGAGCT 1889
|||||
Db 398 AGTTGGCTTCTAATCAAGCTTTTAAACCCCTATTGCTTAAGAAATGGAAGGTGGAGAGCT 339
|||||
QY 1890 CCTGAAGTAAGCAAGACTTTCCTTAGTCGACCCAGTTAAGAAATCTTCTATGTG 1949
|||||
```

(Pharmacia), digested with Pac I and Eco RI sites of the modified p7T3 vector. Library was then through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 143 a 101 c 127 g 185 t 7 others

Query Match 18.9%; Score 421.4; DB 9; Length 563;  
Best Local Similarity 92.2%; Pred. No. 2.5e-98;  
Matches 519; Conservative 0; Mismatches 32; Indels 12; Gaps 7;

QY 1395 AAACGCAACCTGTAATGTTGTTAAAGAGTTAGTTGAGTTGCTATCATGCAACAGTGA 1454  
|||||  
Db 1 AAACGCAACCTGTAATGTTGTTAAAGAGTTAGTTGAGTTGCTATCATGCAACAGTGA 60  
QY 1455 AAATGCTGTATTAGTACACAGAGATAATCTAGCTTTGAGCTTAAGAAATTTTGACGAGTG 1514  
|||||  
Db 61 AAATGCTGTATTAGTACACAGAGATAATCTAGCTTTGAGCTTTAAGAAATTTTGACGAGTG 120  
QY 1515 GTATGTTGGGAGACTGCTGAGTCAACCCCAATAGTTGATTGGCAGGAGTTGGAAGTG 1574  
|||||  
Db 121 GTATGTTGGGAGACTGCTGAGTCAACCCCAATAGTTGATTGGCAGGAGTTGGAAGTG 180  
QY 1575 TGTGATCTGTGGCACATTAAGCTATGTCATGAGCATCTAAGTAATGATGTCGTTGA 1634  
|||||  
Db 181 TGTGATCTGTGGCACATTAAGCTATGTCATGAGCATCTAAGTAATGATGTCGTTGA 240  
QY 1635 ATCAGATATACGCTCCATCGCTGTGTCATCTAGCTGAGTCCATCTCTCAGGCTTGCT 1694  
|||||  
Db 241 ATCAGATATACGCTCCATCGCTGTGTCATCTAGCTGAGTCCATCTCTCAGGCTTGCT 300  
QY 1695 GCCAAAGCC-TTTTCTGTTTCTTTTGTATCATTAATGAAGTCATGCGTTTAATCATT 1753  
|||||  
Db 301 GCCAAAGCC-TTTTCTGTTTCTTTTGTATCATTAATGAAGTCATGCGTTTAATCATT 360  
QY 1754 CGAG--TGTTTCAGTCTTCGAGATGCTCTGATGCTCATATTTGTTCCCTAATTTG-C 1809  
|||||  
Db 361 CGAGTGGTGTTCAGTCTTCAGATGTCCTTGATGCTCATATTTGTTCCCTAATTTGCGC 420  
QY 1810 CAGTGGGACTCCTAA-ATCAATTTGGCTCTAA--TCAAGCTTTTAAACCCCTATTGG 1865  
|||||  
Db 421 AGTTGGGAACCTCCTAAATCAAGTTGGCTCTAAATCCAAGCTTTTAAACCCCTATTGG 480  
QY 1866 TAAAGAAATGG-AAGTGGGAGAGCTCCCTGAAAGTAAAGCAAGAC--TTTCCCTTTAGTCG 1922  
|||||  
Db 481 TAAAGAAATGGTAAAGTGGGAGAGCTCCCGAGTAATCAAGACCTTTCCNCCCTTAGTCG 540  
QY 1923 ACCCAAGTTAAGATGTTCTTAT 1945  
|||||  
Db 541 AGCCAGGTTNAGATGTTCTNTAT 563

RESULT 5  
AA807940/c  
LOCUS  
DEFINITION nu90f10.s1 NCI\_CGAP\_Pr22 Homo sapiens cdna clone IMAGE:121795 3',  
mRNA sequence.  
ACCESSION AA807940  
VERSION AA807940.1 GI:2877346  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 398)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
www-bio.llnl.gov/hbrp/image/image.html  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 385.

FEATURES  
source

1..398  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:121795"  
/clone\_lib="NCI\_CGAP\_Pr22"  
/sex="male"  
/tissue\_type="normal prostate"  
/lab\_host="DH10B"  
/note="Organ: prostate; Vector: p7T3D-Pac (Pharmacia)  
with a modified polylinker; 1st strand cDNA was prepared  
from normal prostate bulk tissue, and was then primed with  
a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified p7T3 vector. Library is normalized, and was  
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 117 a 87 c 73 g 121 t

Query Match 17.7%; Score 394.8; DB 9; Length 398;  
Best Local Similarity 99.5%; Pred. No. 1.8e-91;  
Matches 396; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1786 TGCTCATATTTGTTCCCTAATTTGCGAGTGGGAAGTCTTAATCAAAATGGCTTTCTAATCA 1845  
|||||  
Db 398 TGCTCATATTTGTTCCCTAATTTGCGAGTGGGAAGTCTTAATCAAAATGGCTTTCTAATCA 339  
QY 1846 AAGCTTTTAAACCCCTATTGTTAAAGAAATGGAAGTGGAGAGTCCCTGAAGTAAGCAAA 1905  
|||||  
Db 338 AAGCTTTTAAACCCCTATTGTTAAAGAAATGGAAGTGGAGAGTCCCTGAAGTAAGCAAA 279  
QY 1906 GACTTTCTCTTAGTCGAGCCAAAGTAAGAATGTTCTTATGTTGCCAGTGTGTTCTGA 1965  
|||||  
Db 278 GACTTTCTCTTAGTCGAGCCAAAGTAAGAATGTTCTTATGTTGCCAGTGTGTTCTGA 219  
QY 1966 TCTGATGCAAGCAAGAAACACTGGGCTTCTAGAACCCAGCAACTTGGGAACTAGACTCCC 2025  
|||||  
Db 218 TCTGATGCAAGCAAGAAACACTGGGCTTCTAGAACCCAGCAACTTGGGAACTAGACTCCC 159  
QY 2026 AAGCTGGACTATGCTCTACTTTTTCAGGCCACATGGCTAAAGAGGTTTCAGAAAGAGTG 2085  
|||||  
Db 158 AAGCTGGACTATGCTCTACTTTTTCAGGCCACATGGCTAAAGAGGTTTCAGAAAGAGTG 99  
QY 2086 GGGACAGCAGCAAGTCTTACCTTCATATATTGTTATGATCTCTAATGATGATGATAAATG 2145  
|||||  
Db 98 GGGACAGCAGCAAGTCTTACCTTCATATATTGTTATGATCTCTAATGATGATGATAAATG 39  
QY 2146 TTAAGTTGATGGTGATGAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 2183  
|||||  
Db 38 TTAAGTTGATGGTGATGAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1

RESULT 6  
BQ027284/c  
LOCUS  
DEFINITION UI-H-C00-ago-h-07-0-UI.s1 NCI\_CGAP\_Sub9 Homo sapiens cdna clone  
IMAGE:3104867 3', mRNA sequence.  
ACCESSION BQ027284  
VERSION BQ027284.1 GI:19762563  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1 (bases 1 to 407)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA-Yes.

FEATURES             Location/Qualifiers
     source
     1..407
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="IMAGE:3104867"
     /clone_lib="NCI-CGAP_Sub9"
     /tissue_type="mixed"
     /dev_stage="mixed"
     /lab_host="DH10B (Life Technologies)"
     /note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoR.I; Site_2: Not I; tissues:
Cholonic mucosa with Crohns disease, Cholonic mucosa with
ulcerative colitis, Fetal thymus, Cervix, Cervical
adenosquamous carcinoma, Ligament cells, Prostate
carcinoma, Bladder carcinoma, Brain oligodendrocyte ;
NCI-CGAP_Sub9 is a subtracted cDNA library constructed
according to Bonaldi, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are CGTC, AACG, GGGCC,
GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional
information, contact: Bento Soares, bento-soares@uiowa.edu
TAG_LIB=UI-H-COO
TAG_ISSUE=Cervical Adenosquamous Carcinoma
TAG_SEQ=CGAAG"

BASE COUNT  115 a  83 c  64 g  145 t
ORIGIN
Query Match      17.5%; Score 390.2; DB 14; Length 407;
Best Local Similarity 99.2%; Pred. No. 2.8e-90;
Matches 392; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  1837 TTCTAATCAAGCTTTTAAACCTTATGGTAAAGAAATGGAGTGGAGAGCTCCCTGAA 1896
Db  407 TTTTAAATCAAGCTTTTAAACCTTATGGTAAAGAAATGGAGTGGAGAGCTCCCTGAA 348
QY  1897 GTAGCAAGAGCTTCTCTTAGTCGAGCCAAAGTAAAGATGTTCTTATGTTGCCCATG 1956
Db  347 GTAAGCAAGAGCTTCTCTTAGTCGAGCCAAAGTAAAGATGTTCTTATGTTGCCCATG 288
QY  1957 TGTCTCTGATCGATGCAAGCAAGAAACACTGGGGCTTCTAGAACACGCAACTTGGGAAAC 2016
Db  287 TGTCTCTGATCGATGCAAGCAAGAAACACTGGGGCTTCTAGAACACGCAACTTGGGAAAC 228
QY  2017 TAGACTCCCAAGCTGACTATGGCTCTACTTTTCAGCCACATGGCTTAAGAGGTTTCAG 2076
Db  227 TAGACTCCCAAGCTGACTATGGCTCTACTTTTCAGCCACATGGCTTAAGAGGTTTCAG 168
QY  2077 AAAGAAGTGGGACAGACGACAACTTTTCACTTATATATTTGATGATCTCTTAATGAATG 2136
Db  167 AAAGAAGTGGGACAGACGACAACTTTTCACTTATATTTGATGATCTCTTAATGAATG 108
QY  2137 CATAAATGTTAAGTTGATGTAATGAAATGTAATGTAATGTTTAAACAACTATGTTG 2196

```

```

||||| CATAAATGTTAAGTTGATGTAATGTAATGTAATGTAATGTTTAAACAACTATGTTG 48
Db  107 GAAATAAATCAAGCTGCTATTAAGTATGTTGATGTTGATGTTGATGTTGATGTTG 2231
QY  2197 GAAATAAATCAAGCTGCTATTAAGTATGTTGATGTTGATGTTGATGTTGATGTTG 13
Db  47 GAAATAAATCAAGCTGCTATTAAGTATGTTGATGTTGATGTTGATGTTGATGTTG 13

RESULT 7
LOCUS H58254/c
DEFINITION Yr06h06.sl Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:204539 3', mRNA sequence.
ACCESSION H58254
VERSION H58254.1 GI:1011086
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 467)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1025
High quality sequence stops: 344
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1025 Std Error: 0.00
Seq primer: Promega -2lml3
High quality sequence stop: 344.
FEATURES             Location/Qualifiers
     source
     1..467
     /organism="Homo sapiens"
     /db_xref="GDB:373670"
     /db_xref="taxon:9606"
     /clone="IMAGE:204539"
     /clone_lib="Soares fetal liver spleen INFLS"
     /sex="male"
     /dev_stage="20 week-post conception fetus"
     /lab_host="DH10B (ampicillin resistant)"
     /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACTGGAAGAATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldi."

BASE COUNT  125 a  98 c  91 g  146 t
ORIGIN
Query Match      16.1%; Score 360; DB 14; Length 467;
Best Local Similarity 96.1%; Pred. No. 2.1e-82;
Matches 421; Conservative 0; Mismatches 12; Indels 5; Gaps 5;

QY  1799 CCTTAATTTGCCAGTGGGAACCTCTAAATCAAAATGGG-CCTTAAATCAAGCTTTTAA- 1856
Db  445 CCTAATTTGCCAGTGGGAACCCCTAAATCAAGTGGCCCTCTTAATCAAGCTTTTAAAC 386
QY  1857 CCCTATTGTTGAAGAAT-GGAAGGTGGAGAGCTCCCTCAAGTAAGCAAGACTTTCCTC 1915

```

Db 385 CCTATGGTAAAGTAAGGAGGAGGAGAGCTCCTCGAAGTAAGCAAAAGACTTTCTC 326

QY 1916 TTATGTCGAG-CCAAGTTAAGAAATGTTCTTATGTTGCCAGTGTGTTCTCATCTGATGCA 1974

Db 325 TTATGTCGAGCCCAAGTAAAGATGTTCTTATGTTGCCAGTGTGTTCTCATCTGATGCA 266

QY 1975 ACAGAAACACTGGGCTTCTAGAA-CCAGCAACTTGGGAACCTAGACTCCCAAGCTGGA 2033

Db 265 AGCAAGAAACACTGGGCTTCTAGAAACCCAGCAACTTGGGAACCTAGACTCCCAAGCTGGA 206

QY 2034 CTATGGCTCTACTTTCAGGCCACATGGCTAAAGAGAGTTTCAGAAAGAGTGGGGACAGA 2093

Db 205 CTATGGCTCTACTTTCAGGCCACATGGCTAAAGAGAGTTTCAGAAAGAGTGGGGACAGA 146

QY 2094 GCAGAACTTTCACCTTCATATATTGTTATGATCCCTAAATGAATGCAATAAATCTTAAGTGG 2153

Db 145 GCAGAACTTTCACCTTCATATATTGTTATGATCCCTAAATGAATGCAATAAATCTTAAGTGG 86

QY 2154 ATGTTGATGAATCTAAATCTGTTTAAACAATATGATTTGGAAAAATAAATCAATGCT 2213

Db 85 ATGTTGATGAATCTAAATCTGTTTAAACAATATGATTTGGAAAAATAAATCAATGCT 26

QY 2214 ATRACTATGTTGATAAAA 2231

Db 25 ATRACTATGTTGATAAAA 8

RESULT 8

LOCUS BE656336

DEFINITION UI-M-BHO-aju-h-06-0-UI.r1 NIH\_BMAP\_M.S1 Mus musculus CDNA clone

ACCESSION UI-M-BHO-aju-h-06-0-UI 5', mRNA sequence.

VERSION BE656336

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

COMMENT Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov  
CDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Researchers may obtain BMAP CDNA clones from RESEARCH GENETICS. It  
should be noted that Bento Soares is generating a small number of  
additional specialized non-redundant arrays of BMAP cDNAs whose  
availability will be considered under appropriate and limited  
collaborative arrangements  
Seq primer: M13 Reverse.

FEATURES

source

1. 487

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-BHO-aju-h-06-0-UI"

/clone\_lib="NIH\_BMAP\_M.S1"

/dev\_stage="27-32 days"

/lab\_host="DH10B (Life Technologies)"

/notes="vector: pRT3D-Pac (Pharmacia) with a modified  
polylinker; Site.1: Not I; Site.2: Eco RI; The  
NIH\_BMAP\_M.S1 library is a subtracted library derived from  
a mixture of normalized libraries from ten regions of the  
mouse brain (cerebellum, brain stems, olfactory bulbs,

hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 20,000 cDNA clones obtained from non-normalized and normalized libraries of these ten regions of the mouse brain."

BASE COUNT 115 a 115 c 103 g 154 t

Query Match 16.0%; Score 357.6; DB 10; Length 487;  
Best Local Similarity 85.3%; Pred. No. 8.9e-82;  
Matches 399; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 110 TTGATTATGATTAGCGTCCCTGTCTATAAATTTGACGTGAAGCAAAATTTGGGGCCCAAC 169

Db 19 TCGATTATGTTATGTCAGCACCCCTGCCAAAAAATCAATGTGAACAAATTTGGGCTCAGC 78

QY 170 TCCTGCCCTCCGCTCTACTCGCTGTTTCATCTTTTGGTTTGTGGGCAACATGCTGGTGG 229

Db 79 TCCTGCCCTCCGCTCTACTCGCTGTTTCATCTTTTGGTTTGTGGGCAACATGCTGGTGG 138

QY 230 TCCTCATCTTAAATAAAGCTGCAAAAGCTGAAGTCTTTCACCTGACATATTTACCTGCTCAACC 289

Db 139 TCCTCATCTTAAATAAAGCTGCAAAAGCTGAAGTCTTTCACCTGACATATTTACCTGCTCAACC 198

QY 290 TGGCCATCTCTGATCTGCTTTTCTTATTACTCTCCCATTTGGGGCTCACTGCTGTCGAA 349

Db 199 TGGCCATCTCTGATCTGCTTTTCTTATTACTCTCCCATTTGGGGCTCACTGCTGTCGAA 258

QY 350 ATGAGTGGGCTTTTGGGAATGCAATGTGCAAAATTTATTCACAGGCTCTATCACATCGTT 409

Db 259 ATGAGTGGGCTTTTGGGAATGCAATGTGCAAAATTTATTCACAGGCTCTATCACATCGTT 318

QY 410 ATTTGGGGGAATCTTCTTCATCATCTCCTCTGACATCGATAGATACCTGGCTATTGTC 469

Db 319 ATTTGGGGGAATCTTCTTCATCATCTCCTCTGACATCGATAGATACCTGGCTATTGTC 378

QY 470 ATGCTGTGTTTCTTAAAGCCAGGACGGTCACTTTCGGGGTGTGACAAAGTGTGATCA 529

Db 379 ATGCTGTGTTTCTTAAAGCCAGGACGGTCACTTTCGGGGTGTGACAAAGTGTGATCA 438

QY 530 CCTGTTGGTGGCTGTGTTTGTCTTGTCTCCAGGAATCATCTTTACTA 577

Db 439 CTTGGGGTGGCTGTGTTTGTCTTGTCTCCAGGAATCATCTTTACTA 486

H58245 461 bp mRNA linear EST 05-OCT-1995  
IMAGE:204515 3', mRNA sequence.

ACCESSION H58245

VERSION H58245.1 GI:1011077

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 461)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, J., Waterston

, R., Williamson, A., Wohldmann, P. and Wilson, R.

Unpublished (1995)

Contact: Willson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1007

High quality sequence stops: 335

Source: IMAGE Consortium, LLNL









**ORGANISM**  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**REFERENCE**  
 1 (bases 1 to 447)  
**AUTHORS**  
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
 Thelsing,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
 Waterston,R.  
**TITLE**  
 The WashU-RHMI Mouse EST Project  
**JOURNAL**  
 Unpublished (1996)  
**COMMENT**  
 Contact: Marra M/Mouse EST Project  
 WashU-RHMI Mouse EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LML; contact the  
 IMAGE Consortium (info@image.lml.gov) for further information.  
 MGI:552008  
 Seq primer: -28ml3 rev2 ET from Amersham  
 High quality sequence stop: 397.  
**FEATURES**  
 Location/Qualifiers  
 1..447  
   /organism="Mus musculus"  
   /strain="C57BL/6J"  
   /db\_xref="taxon:10090"  
   /clone="IMAGE:963216"  
   /clone\_lib="Soares\_mammary\_gland\_NBMMG"  
   /sex="male"  
   /tissue\_type="mammary gland"  
   /dev\_stage="4 weeks"  
   /lab\_host="DH10B"  
   /note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia  
   ) with a modified polylinker; Site\_1: Not I; Site\_2: Eco  
   RI; 1st strand cDNA was primed with a Not I - oligo(dT)  
   primer [5'  
   TGTACCAATCTGAAGTGGGCGCGCGAATGGTTTCTTTTCTTTTCTTTTCTTTT  
   T 3']; double-stranded cDNA was ligated to Eco RI  
   adaptors (Pharmacia), digested with Not I and cloned into  
   the Not I and Eco RI sites of the modified pT73 vector.  
   RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
   constructed and normalized by Bento Soares and M.Patima  
   Bonaldo."  
**BASE COUNT** 104 a 102 c 92 g 149 t  
**ORIGIN**  
 Query Match 13.1%; Score 292.6; DB 9; Length 447;  
 Best Local Similarity 80.7%; Pred. No. 6.7e-65;  
 Matches 359; Conservative 0; Mismatches 74; Indels 12; Gaps 1;  
 QY 383 TATTACAGGGGTGATACATCGGTATTTTGGCGGAATCTTCTTCATCATCTCTCTGA 442  
   |||||  
 Db 2 TATTACAGGGGTCTATCATTTGGTATTTTGGTGAATCTTCTTCATCTCTCTGA 61  
 QY 443 CAATCGATAGATACCTGGCTATTGTCACCTCTGTTTCTTTAAAGCCAGGAGTCA 502  
   |||||  
 Db 62 CAATTGATAGGTACTTGGCTATTGTCATGCTGTTTCTTTAAAGTCTCTCTCTGA 121  
 QY 503 CCTTTGGGTGTGACAAAGTGTATCATCACTGGTTGGTGTGTTGCTTCTCTCCAC 562  
   |||||  
 Db 122 ACTTTGGGGTGATACAAAGTGTAGTCACTTGGCGGTGGCTGTGTTTCTCTCTCCAG 181  
 QY 563 GAATCATCTTTACTAAATGCCAAGAGATCTGTTTATGTCGTGGGCCCTTATTTTC 622  
   |||||  
 Db 182 AAATAACTTTTACCAGATCTCAGAAGAAGGTTTTCATTATATACATCAGTCCTCATTTTC 241  
 QY 623 CA-----CGAGGATGGAAATAATTTCCACACAATAATGAGGAACATTTGGGC 670  
   ||  
 Db 242 CACACACTCATGATCATATTTCTGGAAGAGTTTCCAAACATTAAGATGTCATCTTGAGCC 301  
 QY 671 TGGTCCTGCCGCTGCTCATCATGTCATCTGCTACTCGGGAATCCTGAAACCCCTGCTTC 730

Mon Jun 2 09:42:08 2003

QY 280 CTGCTACCTGGCCATCTCTGATCTGCTTTTCTTATTAATCTCCCAATCTGGGCTCAC 339  
 Db 243 CTCCTGAACCTGGCCATCTCTGATCTGCTTTTCTTATTAATCTCCCAATCTGGGCTCAC 302  
 QY 340 TCTGCTGCAA---ATGAGTGGGTCTTTGGGAATGCAATGTGCAAAATTAATCACAGGGCTG 396  
 Db 303 TACAACTTGAAGGATGACTGGGTCTTTGGTATGCAATGTGTAAGATCTCTCTGGGTTT 362  
 QY 397 TATCACATCGGTATTTTGGGGATCTCTTATTAATCTCCCAATCTGGGCTCAC 456  
 Db 363 TATTACACAGGCTTTACAGCGAGATCTTTTATTAATCTCCCAATCTGGGCTCAC 422  
 QY 457 CTGGCTATTGTCATGCTGTGTTTAAAGCCAGGACGCTCACCTTTGGGGTGGTG 516  
 Db 423 CTGGCCATCGTCCAGCGGTGTTGGCTTGGGGCAGGACGCTCACCTTTGGGTGTCATC 482  
 QY 517 ACAAGTGTGATCACCTGGTGGTGGTGTGTTGCTTCTGTCCTCCAGGATCATCTTACT 576  
 Db 483 ACCAGCATCATCTTTGGGCCCTGGCCATCTTTGGGCTTCCATGCCAGGCTTATACITTC 542  
 QY 577 AATGCCAGAAAGATCTCTTATGCTGTGGCCCTTATTTCC-----A 624  
 Db 543 AAGACCAATGGGAATTTCACTCACACACCTGCAGCCTTCACTTTCTCACGAAAGCCTA 602  
 QY 625 CGAGGATGGAATAATTTCCACACAATAATGAGAAATTTGGGGCTGGTCTGCGGCTG 684  
 Db 603 CGAGAGTGAAGCTGTTTCAGGCTCTGAAACTGAACCTCTTTGGGGTGGTATGCGCTTG 662  
 QY 685 CTCATCATGTCATCTGCTACTCGGGAATCCTGAAACCTGCTCGGTGCGAAACGAG 744  
 Db 663 TTGGTCATGATCATCTGCTACACAGGATTAAGATTTCTGCTAAGACGACCAATGAG 722  
 QY 745 AAGAAGAGGCATAGGCGCAGTGAGATCATCTTCCACCATCATGATTTTACTTCTCTTC 804  
 Db 723 AAGAA---ATCCAAAGCTGTCCTGTTGATTTTGTATCATGATCATCTTTTCTCTTT 779  
 QY 805 TGGACTCCCTATAACATTTGCTATCTCTGACACACCTTCCAGGAATTTCTCGGCTGAGT 864  
 Db 780 TGGACCCCTTACNATTTGATATATTTCTGTTTCCAGACTTCTCTGTTACCCAT 839  
 QY 865 AACTGTGAAGCAGCAGTCAACTGAGCAAGCCAGCAGGCTGACAGAGACTCTTGG 920  
 Db 840 GAGTGTGAGCAGCAGCAGATTTGGACCTGGCTGTGCAAGTGAACGGAAGTATCG 895

Search completed: June 1, 2003, 20:03:35  
 Job time : 2826.92 secs

